Univerza v Ljubljani
Fakulteta za računalništvo in informatiko

Rok Povšič

Napovedovanje donosa delnic s pomočjo računovodskih podatkov

magistrsko delo

magistrski program druge stopnje
računalništvo in informatika

Mentor: izr. prof. dr. Polona Oblak
Somentor: doc. dr. Erik Štrumbelj

Ljubljana, 2016

University of Ljubljana
Faculty of Computer and Information Science

Rok Povšič

Forecasting stock returns using accounting data

MASTER'S THESIS

SECOND CYCLE STUDIES
COMPUTER AND INFORMATION SCIENCE

Advisor: assoc. prof. dr. Polona Oblak
Co-advisor: assist. prof. dr. Erik Štrumbelj

Ljubljana, 2016

I would like to thank my advisor, assoc. prof. dr. Polona Oblak, and co-advisor assist. prof. dr. Erik Štrumbelj for their guidance. I am especially grateful to Luka Gubo for his expert advice. I would also like to thank Eva and my parents for their support.

Contents

Seznam uporabljenih kratic

kratica	angleško	slovensko
P/E	price to earnings	cena na dobiček
P/B	price to book value	cena na knjigovodsko vrednost
P/S	price to sales	razmerje med ceno in prihodki
EV	enterprise value	vrednost celotnega kapitala
EBIT	earnings before interest and taxes	dobiček iz poslovanja
EPS	earnings per share	dobiček na delnico

Povzetek

Naslov: Napovedovanje donosa delnic s pomočjo računovodskih podatkov

Srednje močna oblika teorije učinkovitih trgov pravi, da so vse javne informacije vsebovane v cenah delnic. Glede na to teorijo javni računovodski podatki ne morejo imeti sposobnosti napovedovanja prihodnjih presežnih donosov. Z uporabo letnih in četrtletnih računovodskih podatkov, ki niso pristranski v korist preživelih, na podjetjih Združenih držav Amerike in Evrope ovrednotimo več naložbenih strategij, ki delujejo na podlagi podcenjenosti, kakovosti in statističnega modeliranja. Ugotovimo, da imajo določene strategije z uporabo teh podatkov neko napovedno sposobnost in da preproste strategije, ki investirajo v podjetja, ki so hkrati kakovostna in podcenjena, delujejo najbolje.

Ključne besede: delnice, donosi, napovedovanje, računovodski podatki.

Abstract

Title: Forecasting stock returns using accounting data

Semi-strong form of the Efficient Market Hypothesis states that all public information is reflected in stock prices. According to this theory, publicly available accounting data should not have the ability to predict future relative stock returns. We take annual and quarterly survivorship bias-free accounting data for companies in the United States and Europe and test several value, quality, and statistical modeling strategies. We find that certain strategies utilizing this data do have predictive ability and that simple strategies that invests in companies that are both of quality and undervalued, work best.

Keywords: stocks, returns, forecasting, accounting data.

Razširjen povzetek

Borze po svetu so v zadnjih sto letih na dolgi rok v povprečju ustvarjale pozitivne realne donose. Od leta 1900 do 2011 je bila mediana letnih rasti borz 4.6% po upoštevani inflaciji, sama letna inflacija pa 4.1%. Na krajši rok je prihajalo do visokih nihanj, nazadnje predvsem ob finančni krizi v letih od 2007 do 2009, ko so vrednosti delnic borz po svetu v povprečju izgubile tudi po 50%. Še hujši padec se je zgodil v tridesetih letih prejšnjega stoletja, ko so vrednosti ameriških delnic v povprečju padle za več kot 80%. Tveganje, ki ga investitor prevzame za dolgoročne donose je, da se na kratki rok lahko zgodijo veliki padci. Toda za padci so sledila okrevanja in dolgoročni investitorji so dosegali pozitivne povprečne donose.

Z nakupom delnice investitor postane delni lastnik podjetja. Nakupe in prodaje lahko izvršuje preko borznega posrednika. To je bila v preteklosti fizična oseba, v času interneta pa se čedalje več uporablja spletne borzne posrednike, zaradi česar so se transakcijski stroški nakupa in prodaje znižali.

Neposredni nakup delnic za povprečnega investitorja vseeno predstavlja določen izziv, saj mora izbrati primerna podjetja, naložbe razpršiti v dovolj veliko število podjetij, posodabljati portfelj, ipd. Zaradi tega so se pojavili vzajameni skladi, ki te stvari počnejo namesto investitorja. Vzajemni sklad ima tipično v lasti veliko število podjetij in zaradi te razpršitve se tveganje celotnega portfelja zmanjša. Transakcijski stroški se razdelijo na več oseb, zato v povprečju investitor za njih plača manj kot pri lastnem upravljanju. Storitev upravljanja premoženja prek vzajemnih skladov tipično letno stane od 1% do 2% upravljanega premoženja.

Alternativa vzajemnim skladom so skladi, ki kotirajo na borzi (angl. Exchange Traded Fund oz. ETF). Njihova prednost je ta, da so stroški upravljanja nižji, kar ima na dolgi rok velik vpliv.

Doseganje povprečnih donosov trga delnic je enostavno. Z nakupom nizkostroškovnega ETF-ja, kot je npr. ameriški SPY ali nemški DAX, se investitor izpostavi do celotnega trga in dosega povprečne donose. Toda doseganje nadpovprečnih donose je težje, kar se vidi že iz dejstva, da večina vzajemnih skladov ne dosega niti povprečnih donosov trga, v katerega so investirani.

A vendar je poskus doseganja nadpovprečnih donosov lahko vreden svojega truda, saj ima na dolgi rok že relativno majhna izboljšava donosov velik vpliv. Če uspemo povprečni donos investicijske strategije izboljšati z 8% na 9%, bomo v 30 letih iz začetne investicije \$10,000 namesto \$100,627 dosegli \$132,677.

V magistrskem delu z uporabo letnih in četrtletnih računovodskih podatkov največjih 500 ameriških in največjih 500 evropskih podjetij simuliramo različne investicijske strategije z namenom, da najdemo tiste, ki izboljšajo povprečne donose. Glede na teorijo učinkovitih trgov naj to ne bi bilo mogoče, saj se vsaka javna informacija v trenutku objave že odraža v ceni delnice, zaradi česar računovodskih podatkov ne moremo uporabiti za konsistentno doseganje nadpovprečnih donosov.

Najprej analiziramo enostavne investicijske strategije, ki poskušajo najti podjetja, ki so t.i. *podcenjena*. Natančna definicija *podcenjenosti* ne obstaja, uporabimo pa lahko eno izmed razmerij, ki jo poskuša zajeti. Eno izmed tradicionalnih razmerij je razmerje P/B oz. razmerje med ceno delnice in knjigovodsko vrednostjo podjetja na delnico. Tako razmerje najprej uporabimo v strategiji, kjer enkrat na leto izberemo podjetja, ki so na voljo v zadnjem letu, in jih razdelimo na kvintile. Izračunamo aritmetično povprečje donosov vseh delnic posameznega kvintila v nekem letu ter nato geometrijsko povprečje za vsak kvintil skozi vsa leta. S tem dobimo povprečen donos, ki ga lahko primerjamo z ostalimi kvintili in ocenimo, ali ima razmerje sposobnost ločiti podjetja glede na njihove prihodnje donose.

Donose strategij preverjamo na borznem trgu Združenih držav Amerike in, v primeru, da strategija deluje, še na evropskem trgu. Dostop do računovodskih podatkov nekega fiskalnega leta ali fiskalnega četrtletja podjetja dovolimo po relativno dolgem časovnem zamiku, da se izognemo temu, da bi v simulaciji na neki točki uporabili informacije, ki še niso bile javno dostopne. Pri določenih simulacijah uporabimo tudi točne podatke o tem, kdaj so bili računovodski podatki objavljeni.

Pri strategijah, ki uporabljajo enostavna razmerja podcenjenosti, smo ugotovili, da sta izmed testiranih razmerij imela najboljšo napovedno sposobnost razmerje med ceno delnice in prihodki na delnico ter razmerje med vrednostjo celotnega kapitala in dobičkom iz poslovanja. Z uporabo teh strategij lahko z investicijo v delnice najbolj podcenjenega kvintila povprečne donose trga izboljšamo za 1,86% oz. 1.90%.

Naslednja strategija, ki smo jo testirali, ne poskuša iz računovodskih podatkov ugotoviti podcenjenosti, temveč kakovost podjetja. To stori tako, da za vsako letno poročilo podjetja, ki vsebuje vse potrebne podatke, izračuna 9 binarnih vrednosti. Te ocenjujejo lastnosti podjetja, kot so profitabilnost, finančni vzvod, likvidnost, ekonomičnost ipd. Končni indikator je vsota teh 9 vrednosti. Visoke vrednosti končnega indikatorja naj bi kazale na to, da je podjetje kakovostno. Simulirali smo zgodovinske nakupe delnic in ugotovili, da podjetja z visokim indikatorjem posledično prinašajo višje donose, kot tista z nizkim. Povprečni donosi podjetij, ki imajo določeno vrednost indikatorja, so skoraj popolnoma monotoni - nižja kot je končna vrednost indikatorja, nižji so donosi. Ugotovili smo, da investicije v delnice z indikatorjem 7, 8 ali 9 na ameriškem trgu v povprečju izboljšajo letne donose za 1.45%. Ob posodabljanju portfelja štirikrat na leto je investitor na ameriškem trgu tako lahko dosegal geometrijsko povprečne donose 10,60%, na evropskem pa 10,88%. Z uporabo te strategije smo na ameriškem trgu lažje odkrili podjetja, ki bodo v prihodnje imela nižje donose, na evropskem trgu pa podjetja s prihodnjimi višjimi donosi.

Naslednji dve strategiji, ki smo ju analizirali, uporabljata statistično učenje na veliki količini računovodskih podatkov. Pri obeh uporabljamo okno dolžine 10 let, na katerem gradimo model in ga ovrednotimo na letu, ki sledi. Nato okno premaknemo za leto naprej in ponovimo, vse do zadnjega leta. Na tak način z uporabo podatkov od leta 1995 do 2013 strategiji testiramo na vseh letih od 2005 do 2013.

Prva strategija z uporabo logistične regresije modelira prihodnjo smer gibanja dobičkov. Iz velike količine računovodskih podatkov najprej izbere le tiste, ki imajo že sami po sebi neko napovedno spodobnost. Količino različnih računovodskih podatkov želimo omejiti, saj lahko vrednosti manjkajo, model pa zahteva, da imajo vsi izbrani podatki vrednost, sicer se računovodski izkaz za neko podjetje v nekem obdobju ne uporabi. Z omejitvijo količine računovodskih podatkov lahko posledično uporabimo več izkazov, kot bi jih sicer, kar pomeni, da imamo na voljo več instanc za učenje in ovredonotenje modela. Izbrane podatke uporabimo v logističnem modelu. Odvisna spremenljivka je, v primeru te strategije, razlika med dobičkom naslednjega leta in dobičkom trenutnega leta, kjer odštejemo trend razlik v zadnjih 4 letih. Ugotavljamo, da bi v teoretičnem primeru, ko bi imeli informacijo o prihodnji smeri spremembe dobička, lahko izboljšali donose trga. Z uporabo logističnega modela, ki vrača verjetnost, da bo prihodnja sprememba dobička pozitivna ali negativna, pa napovedi na naših podatkih niso bile dovolj dobre, da bi zanesljivo premagale povprečne donose trga.

Druga strategija predpostavlja, da imajo podjetja v povprečju učinkovito tržno kapitalizacijo, ki je definirana kot cena delnice pomnožena s številom izdanih delnic, se pa lahko pri posameznih podjetjih pojavijo razlike med tržno kapitalizacijo in dejansko notranjo vrednostjo podjetja. Z uporabo linearne regresije se strategija poskuša naučiti, kaj naj bi bila primerna tržna kapitalizacija nekega podjetja. To počne na podlagi 14 računovodskih podatkov iz zadnje četrtletne bilance stanja in 14 iz vsote zadnjih štirih četrtletnih izkazov poslovnega izida. Podobno kot pri prejšnji strategiji se tudi tu zahteva, da so na voljo vsi podatki, drugače se podjetja v nekem četrtletju ne uporabi. V testnem delu podatkov strategija na podlagi modela vsa podjetja razdeli v kvintile in simulira gibanje vrednosti portfelja vsakega kvintila. Strategija ima določeno sposobnost ločevanja podjetij, ki bodo imela višje donose od povprečja, vendar ta sposobnost ni tako dobra, kot jo vidimo pri ostalih strategijah.

Pri obeh strategijah, ki uporabljata statistično učenje, smo odkrili le omejene izboljšave donosov. Razlog za to je morda ta, da ti dve strategiji uporabljata veliko količino podatkov in ne le tiste, ki imajo intuitiven ekonomski smisel.

Na koncu smo združili dve strategiji - eno, ki išče podcenjena podjetja in drugo, ki išče kakovostna podjetja. Ugotovili smo, da se v tem primeru donosi še nekoliko izboljšajo, spremembe pa so v primerjavi z uporabo samo strategije iskanja podcenjenosti podjetja relativno majhne.

V delu smo ugotovili, da od strategij, ki smo jih obdelali, enostavne delujejo boljše. Prednost imajo strategije, ki imajo za svojo izbiro računovodskih vrednosti nekakšno intuitivno ekonomsko logiko. Modela, ki za izbiro vrednosti uporabljata statistično učenje, se na našem naboru podatkov nista obnašala tako, kot v izvornih člankih.

V primeru nadaljnjega dela bi bilo smiselno strategije izboljšati, da uporabljajo ali modele, ki podpirajo manjkajoče vrednosti, ali pa poskušajo manjkajoče vrednosti zapolniti, npr. z zadnjo znano vrednostjo. Mogoča je tudi optimizacija strategij s tem, da so bolj odzivne in pozicije zajemajo že na točen dan objave računovodskih podatkov. Prav tako bi lahko bila koristna analiza večjega števila podjetij ter delitev podjetij na sektorje in industrije.

Chapter 1

Introduction

This thesis attempts to discover active investment strategies that in the long term achieve better returns than the buy and hold strategy, a strategy that has the whole market in its portfolio. Trading such strategies helps to move stock prices towards their intrinsic value for which financial markets reward us by giving us better returns. By doing that, society as a whole also benefits, because resources can be allocated more fairly. Financial markets are an active research subject as every discovery of an anomaly leads to a better financial market understanding.

This work will attempt to improve returns of the buy and hold strategy by using companies' yearly and quarterly accounting data (also called fundamental data) that are publicly available. We will look at different types of strategies that use accounting data to predict future returns and we will use the same test methodology, so we can compare them.

We are interested in finding out if we are able to replicate on our dataset the results of strategies that were successful in the past and published in academic journals.

1.1 Related work

Basu (1977) has investigated the ability of the Price to Earnings ratio to forecast future returns. He concluded that subsequent stock returns for companies with low Price to Earnings ratio tend to be higher than for companies with high ratio, meaning that the ratio is not fully reflected in stock prices.

Jensen et al. (1997) have similarly analyzed how Price to Book value ratio and company size correlate with subsequent stock returns. The conclusion was that there is a significant premium for small companies and companies with low Price to Book ratios but only in expansive monetary policy periods, i.e. when central banks are increasing the money supply. The excess returns are in some instances negative when the monetary policy is restrictive.

Piotroski (2000) assigns each company a score from 0 to 9 by considering what the company's factors, such as profitability, leverage, and operating efficiency, are. The score is supposed to represent the quality of a company and he observes that companies with a high score tend to perform better in the subsequent 12 and 24-month period.

Ou and Penman (1989) utilize statistical learning by taking annual accounting statements and building a logistic regression model from a large number of most commonly reported items. The model is trained to predict a binary value that is the next year's change in earnings where trend is accounted for. They find improved returns for a long and short portfolio.

Bartram and Grinblatt (2015) apply fundamental analysis with cross-sectional regression using data from quarterly financial statements as predictors. They build a linear model out of a large number of most common accounting items and try to model market capitalization. Their assumption is that, on average, companies are fairly valued and that if a company's predicted market capitalization on test data differs from actual market capitalization, a long or short opportunity is possible.

Bradshaw et al. (2006) analyze and document negative correlation between net external financing and future stock returns. Mohanram (2005) develops an indicator GSCORE that combines traditional fundamental items and ratios with additional ratios measured for growth companies. A strategy using this indicator results in excess returns for long and short side but where most excess returns are generated on the short side. Setiono and Strong (1998) apply fundamental analysis on the prediction of United Kingdom stock market and find that fundamental information could be used in earning abnormal returns. Likewise, Alexakis et al. (2010) find similar results on the emerging Greek stock market.

The remainder of the thesis is organized as follows. In Chapter 2 we provide an overview of the financial markets and investing. We explain what kind of financial instruments exist and their typical costs. An explanation of what it means to beat the stock market is provided and how one might approach doing that. In Chapter 3 we run backtests for several commonly used ratios that try to answer whether a company's stock price is too high or too low. In Chapter 4 we analyze more complex strategies from published papers that utilize different strategies in trying to improve returns. In Chapter 5, two best strategies that capture different type of companies are combined in an attempt to improve the results even further. Finally, in Chapter 6, we present our conclusions about the work and provide directions for further work.

Chapter 2

An overview of stock market

In this chapter we provide a brief overview of how stock market investing works, and different types of brokerages. We explain what mutual funds are, what their costs are, and show the long-term impact of the costs. We also provide a cheaper alternative to mutual funds, explain what it means to beat the market, and how one might try to do that.

2.1 Investing in the stock market

Over the past hundred or so years the stock markets over the world have on average been rising. An investor that has been invested in the markets with a well diversified portfolio for a timeframe of a few decades has almost surely multiplied the investments. From 1900 to 2011 a median average annual return has been 4.6% after inflation, while median inflation has been 4.1% (Faber (2007)). In the long run, the stock market beats not only inflation but additionally compensates the investor.

Of course, by investing in the stock market, the investor becomes exposed to price fluctuations of the stock prices, both rises and falls. In the short term anything can happen as we have seen with the 2007-2009 financial crisis, where stock markets in developed world lost around 50% of their value. Even worse was the stock market crash in the United States in the 1930's, where the market went down for more than 80%. The risk the investor takes is that in exchange for long-term returns the short-term value can suffer.

What does it mean to buy a stock in a stock market? Buying a stock essentially means buying a share of the ownership of a company. That ownership entitles you to getting a part of the dividends that could be paid out by the company. It also makes it possible for you to sell your share at some later time in order to profit on the possible rise in company price. Of course, by owning a share of a company you also risk that no dividends will be paid and that the company's price will decrease.

2.2 Traditional and online brokerages

How to go about buying stocks? Traditionally, the investor opened a broker account, most likely with a local broker. The broker then bought and sold stocks based on the investor's instructions.

With the rise of internet, online brokerages became available. With online brokerages you do not need another person, the broker, to do the trades for you. With a computer and an internet connection you can do all the buying and selling yourself. Online brokerages have several advantages:

- lower costs than in traditional brokerages,
- instant access online,
- complete transparency,
- ability to write software for automated trading.

The downside of an online brokerage is that nobody will warn you about the potential dangers in the market or the software.

At a typical brokerage, some transaction costs are charged when buying or selling a stock (or some other financial instrument). Most academic papers do not apply these transaction costs to their models' performance. In our research, we do account for conservative transaction costs that a typical individual investor would have to pay in the period analyzed.

2.3 Mutual funds

For an average investor there is a downside to buying stocks directly. First, which stocks should he or she even buy? How to select companies that have a good chance that their stock price will rise? How to make sure that not all investment money is lost and we go bankrupt? Second, if an investor chooses to put all of his investment money in a small number of companies there is a danger that some of those companies' prices will decrease a lot and the investor's investment will as well. To reduce this risk, it is recommended to diversify the porfolio by holding a large number of stocks. A well-diversified portfolio requires a lot of stocks and because each transaction costs a certain amount of money, it might not be economical for the investor to manage the porfolio as these costs add up and can substantially hurt the value of the portfolio.

That is why investing in a mutual fund might be better. The mutual fund industry has provided to the general public the ability to be invested in the stock market by handling the work of selecting stocks and doing the actual buying and selling. By holding a larger number of stocks, mutual funds also benefit from diversification by reducing the risk of a single or a few companies losing a lot of value. Since a larger amount of money is being handled, transaction costs are spread across multiple investors.

A person who invests in a mutual fund knows what sectors and what type of financial instruments the mutual fund will be buying, but does not have the burden of doing it all by himself. For their services, mutual funds typically charge from 1 to 2 percent of the managed capital per year.

2.4 Alternatives to mutual funds

An alternative to a mutual fund is an Exchange Traded Fund (ETF). Unlike mutual funds, where fund managers try to pick stocks they think will perform the best, each ETF follows exactly specified rules that the investor knows in advance, on how the money in an ETF will be invested. For example, an ETF might follow one of many stock market indexes. Slight differences between the index and an ETF following it might appear because an ETF also has transaction and management costs. These costs are typically much lower than what a typical investor would otherwise have to pay to get same market exposure.

One of the most popular indexes is Standard & Poor's 500, often abbreviated as S&P 500. It consists of the 500 largest companies that have their common stock listed at the NYSE or NASDAQ stock exchanges in the United States.

The ETF that follows the S&P 500 index is called SPY. Investing money in SPY is basically the same as dividing up your investment money and buying 500 of the largest US companies, where the amount of money each stock gets is proportional to the market capitalization of that company.

Other important indexes are DAX, which consists of 30 major companies on the Frankfurt Stock Exchange, FTSE 100, which consists of 100 largest companies on the London Stock Exchange, and so on.

Why buy an ETF and not the stocks themselves? There are two major reasons. The first is, as we have mentioned before, that buying an ETF ends up costing much less in transaction costs than buying the individual stocks themselves. The second reason is that it is not enough to buy the companies once. Companies grow and shrink and if you want to always have a portfolio consisting of the top 500 companies, you have to sell some companies and buy others. When you hold an ETF, there is no need for doing that.

Same as mutual funds, ETFs typically charge a management fee. However, this fee is usually much lower in ETFs compared to mutual funds, since the rules are known in advance and not much ongoing research is required. If we suppose that an ETF and a mutual fund that invest in the same sector will have equal gross returns in the future, this fee difference makes a large difference in the portfolio value in the long term.

In Figure 2.1 we can see the impact of annual fees. In that period, an equal-weight buy and hold strategy in the largest 500 companies in the United

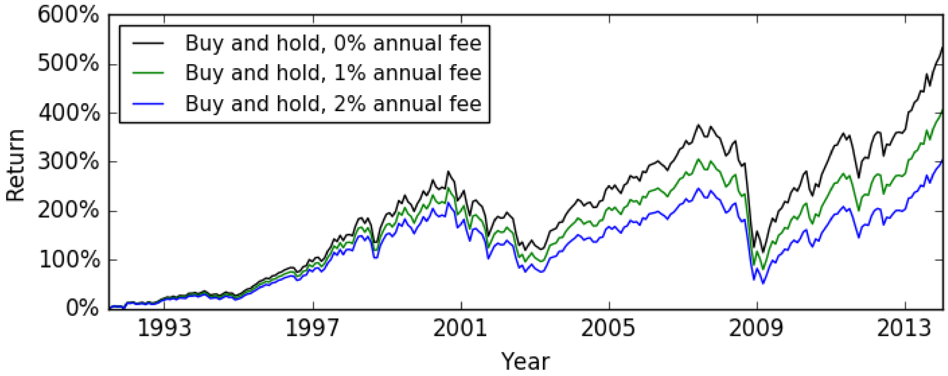

Figure 2.1: Effects of transaction costs on returns for the largest 500 companies in the United States.

States with no fees would turn \$1000 into \$6270, while a mutual fund that performs same as buy and hold would, with a 2% annual fee, result in \$4010.

To make mutual funds attractive for investment, they would have to perform better than ETFs by at least the difference in management fees. However, mutual funds have shown not only to not overperform ETFs by this difference, but even underperform the index on average (Gruber (1996)). Here it is supposed that the mutual fund and the ETF invest in the same market sector.

2.5 Benefits of beating the stock market

If one does not have the ability to pick well-performing stocks, the best bet is to invest in an index. That way the investor is not going to lose the whole investment if a particular company goes bankrupt, which would be the case if the whole portfolio consists of one stock. It is very easy to achieve average market returns. We need only to invest in an index fund or buy an ETF.

Still, the question remains: Is it possible, by picking individual stocks, to get better returns than investing in indexes? When most mutual funds are unable to do that, is there any hope for an average investor?

The reason why most mutual funds underperform is in the fact that few stocks make large returns and the majority make mediocre returns. It is difficult, as a fund or individual stock picker, to pick exactly stocks that will make large returns.

Still, it may be very worth trying to do that, because when investing for a long time, a small increase in the average yearly return makes a big difference. For example, investing $10,000 for 30 years with an 8% average yearly return, will yield approximately $100,627, while a 9% yearly return will yield $132,677, which is a difference of 31.9%.

2.6 Approaches to beating the stock market

One approach to beating the market is analyzing companies in detail by reading reports, watching interviews, talking to CEO, the board, the shareholders, the customers, the competition etc. This approach is time consuming. In this thesis we take another type of approach, a systematic approach, where we analyze a large number of companies exclusively on their accounting data and their stock prices.

There are several systematic approaches one can apply. One approach is buying, for example, 100 stocks that had the best returns in the previous year. Another is buying all the companies in the United States that have their current price above the average price of the last 5 years. These approaches are examples of so-called "technical analysis" strategies because the only inputs into the model are current and past stock prices.

Additionally to stock price data, we will also incorporate into our models accounting data. Strategies that do that are examples of "fundamental analysis" which we use in this thesis.

Usually, a portfolio is updated at regular intervals, because some companies should be sold and other should be bought due to changes prices, fundamental data or some other inputs. These updates are called *rebalances* and might occur yearly, monthly, daily, or at some other time interval.

In this thesis, we will compare all the returns of the strategies to either an equal-weight buy and hold strategy, or to an equal-weight strategy consisting of only those companies that have all the required data. An equal-weight benchmark is used because all our strategies are equally weighted. Comparing our strategies to a market-weight benchmark would mean an advantage in performance since diversified equal-weight portfolios have historically overperformed diversified market-weight portfolios (Bolognesi et al. (2013)), so even a random equal-weight strategy would have an advantage compared to market-weight buy and hold.

Chapter 3

Simple value investing strategies

Before analyzing more complicated models, we look at how simple value investing strategies behaved in the past. Value investing is a strategy where one tries to be invested only in companies that are *undervalued*, expecting the stock price to return to its true intrinsic value in the future.

According to the Efficient Market Hypothesis (Malkiel and Fama (1970)), there should be no such thing as an undervalued company. According to the theory, the price is exactly as it should be and all the publicly available information is already incorporated in the price.[1] From that it follows that it is impossible to predict where the stock price will go in the future even if one has all the public information. However, the assumptions upon which the theory is based, e.g. that all investors are rational and have complete information, do not hold in the real world.

There is not a single definition of what *undervalued* means exactly. Multiple approaches have been developed over the years in trying to capture if a company is undervalued or overvalued. There exist strategies that are based on simple principles and are supposed to tell us something about how much a company is really worth compared to its stock price. These simple strategies usually take some sort of a ratio between what the stock is selling at and what the company's ability to generate value is. The investor then considers

[1]This is the semi-strong Efficient Market Hypothesis.

going long the stocks of the companies that are supposed to be too cheap (i.e. buying them) and *going short* the ones that are supposed to be too expensive (i.e. betting on the price to fall).

We ran backtests for several strategies using such ratios. For each strategy, we split our company universe into several partitions and test performance of each partition to see if there have any been differences in historical returns.

We keep the backtests simple. The steps are as follows:

1. Once per year calculate the ratio for each of the 500 largest companies.

2. Split companies into 5 approximately equally-sized groups (quintiles) based on the ratio, where group Q1 consists of companies that have the lowest ratio value and group Q5 the highest. We only do that for companies with positive ratios and disregard the rest.

3. Calculate the average excess return for each group by subtracting the average return of all the companies from all the groups (i.e. an equal-weight buy and hold) from the average return for each group.

4. Repeat the process for all the years we have data for.

5. Calculate the geometric average of the excess returns. Geometric average instead of arithmetic is here used to account for the fact that a negative return will lose more money that an equal positive will earn.

We do that for all the years between 1991 and 2013, inclusive. All the annual reports can be accessed 5 months after the latest fiscal year end and we assume that they were indeed publicly available at that point. We chose the rebalance date to be July 1, which means we allow access to all the annual reports of fiscal years that ended before February 1 of the current year and after February 1 of the previous year. We chose to do rebalances on July 1 prior to the analysis for the reason that most companies have their fiscal year equal to the calendar year. If we had chosen a later date, the time between accessing the latest annual accounting data would increase for most companies.

The calculation of the geometric average of average annual excess returns is captured in the following equation:

$$\text{Excess return} = \sqrt[n]{\prod_{y=1}^{n} \sum_{j=1}^{N_y} \frac{R_{j,y} - R_{u,y}}{N_y}},$$

where n is the number of years returns are calculated for, $R_{j,y}$ is the annual return of a company j in year y, $R_{m,y}$ is the return of the available universe in year y, and N_y is the number of companies in year y. Available universe consists of only companies that have all the required data. We impose this limitation so we can fairly calculate excess returns.

All excess returns are calculated this way, so strategies can be compared.

3.1 Source of data and software used

In the thesis, we used annual and quarterly accounting data obtained from Bloomberg database (Bloomberg (2016)). Monthly stock price information was used as well and, in certain cases, earnings announcement dates. The prices are adjusted for splits and dividends.

We have written the analyses in Python programming language and used scikit-learn (Pedregosa et al. (2011)) and Statsmodels (Seabold and Perktold (2010)) for statistical learning. Visualization was done with matplotlib (Hunter (2007)).

3.2 Price to Book strategy

The first ratio we analyze is a traditional ratio that is supposed to tell us if a company is undervalued or overvalued. The P/B ratio or Price to Book ratio is defined as:

$$\text{P/B} = \frac{\text{share price}}{\text{book value/shares outstanding}},$$

where *book value* is how much a company's assets are worth according to the accounting statements. Theoretically, this is how much an investor would get if all the assets were liquidated. The rationale behind this ratio is that the larger the ratio, the more we pay for a unit of assets.

We calculated annual excess returns for each group as described before. In Table 3.1 we can see average excess returns compared to buy and hold. The results are not very convincing and we don not find evidence of overperformance with undervalued companies, although the underperformance of the highest quintile is reasonably high, which is what we would expect. However, there should be evidence of overperformance in the bottom quintiles.

Table 3.1: Geometric average of average annual excess returns for Price to Book ratio quintiles.

Quintile	N	Excess return
Q1	1894	0.06%
Q2	1906	-0.30%
Q3	1903	-0.25%
Q4	1906	0.53%
Q5	1914	-2.26%

Book value might not be such an informative accounting item by itself anymore. Companies such as Microsoft or Facebook have a lot of human knowledge and intangible assets that are not necessarily captured in the P/B ratio.

3.3 Price to Earnings strategy

Next, we look at the P/E ratio or Price to Earnings ratio. It is a ratio that gets a lot of coverage in the financial media. Does it have any predictive power?

The P/E ratio is defined as:

$$\text{P/E} = \frac{\text{share price}}{\text{earnings/shares outstanding}}, \tag{3.1}$$

where *earnings* is the last fiscal year's net income or profit. It tries to capture how much we have to pay for a company per unit of earnings.

Table 3.2: Geometric average of average annual excess returns for Price to Earnings ratio.

Quintile	N	Excess return
Q1	1733	0.66%
Q2	1743	-0.05%
Q3	1739	-0.51%
Q4	1743	-0.65%
Q5	1751	-1.05%

In Table 3.2 we can see average excess returns compared to an equal-weight buy and hold portfolio. It seems that the P/E ratio does indeed predict future excess returns as the lowest quintile Q1 is positive, highest quintile Q5 is negative and all the quintiles in between have excess return that falls with each higher quintile.

The number of companies across all the years, as we can see under column N, is lower here than with the P/B ratio, because we disregard all the companies with non-positive ratios, and earnings, as opposed to book value, can be negative. Since price is always positive, the data show that the number of companies with negative earnings is higher than the number of companies with negative book value, which is what we would expect.

3.4 Price to Sales strategy

We now look at the P/S ratio or Price to Sales ratio. It uses sales (also called revenue) instead of earnings. This ratio might be more robust since it is easier for companies to manipulate earnings than sales. The added benefit is that we do not need to exclude companies with negative earnings, since sales is always a positive number.

Table 3.3: Geometric average of average annual excess returns for Price to Sales ratio.

Quintile	N	Excess return
Q1	1912	1.86%
Q2	1922	0.35%
Q3	1918	-0.70%
Q4	1922	-1.45%
Q5	1928	-3.24%

In Table 3.3 we can see the results of the strategy. The results are better compared to using the P/E ratio.

3.5 Enterprise Value to Earnings before Interest & Tax strategy

EV to EBIT ratio or Enterprise Value to Earnings before Interest & Tax ratio is another ratio that tries to capture undervaluation/overvaluation of a company. It is a more accurate ratio of what the true company worth is compared to how much money it is able to generate, where we compare market capitalization with added debt and subtracted cash equivalents, to operating profits.

In Table 3.4 we can see that this ratio is again a good predictor of future excess returns, although not as good as a simple P/S ratio.

3.6 A more detailed analysis

Since P/S ratio worked best, we will analyze it further. We will try improving it by doing a rebalance 4 times a year instead of once a year and using the

Table 3.4: Geometric average of average annual excess returns for Enterprise Value to Earnings before Interest & Tax ratio.

Quintile	N	Excess return
Q1	1594	1.90%
Q2	1602	-0.05%
Q3	1605	-0.01%
Q4	1602	-0.49%
Q5	1612	-2.80%

Table 3.5: Various statistics for P/S ratio quintile portfolios for the 500 largest companies in the United States. Buy and hold (B&H) is included for comparison.

		Price to sales ratio quintile				
	B&H	Q1	Q2	Q3	Q4	Q5
Returns (arithm. avg.)	9.59%	12.56%	11.73%	9.81%	9.34%	8.66%
Returns (geom. avg.)	8.52%	11.19%	10.65%	8.89%	8.11%	5.68%
Volatility	16.46%	19.26%	17.57%	15.70%	17.23%	24.45%
Max Drawdown	-54.87%	-65.04%	-61.95%	-51.56%	-50.12%	-82.28%
Returns / Volatility	0.58	0.65	0.67	0.62	0.54	0.35
$100 becomes	$592	$1,005	$903	$638	$546	$332

data we have on earnings announcement dates to more accurately estimate when we have access to certain accounting data.

Our strategy is the following: every 3 months we take all accounting statements for which we have data that they were released in the past 12 months. We keep this 12-month window in order to have an approximately constant number of companies in our portfolio at all times. If the window size is smaller, large fluctuations appear in the number of companies in different parts of the year since most companies overlap their fiscal year with calendar year.

We can see from Table 3.5 and Figure 3.1 that undervalued quintiles indeed overperform buy and hold. Although Volatility and Maximum Drawdown increase, the average return increases more as seen from improved

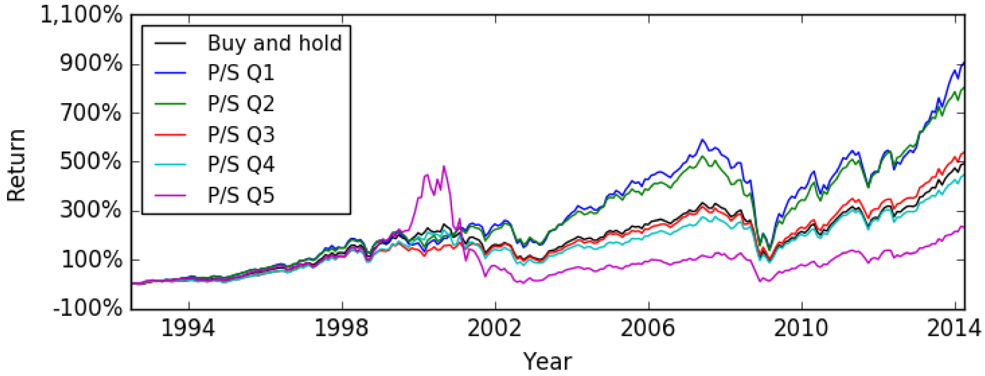

Figure 3.1: Returns for P/S quintiles for the United States market. Buy and hold is included for comparison.

Returns / Volatility.

There are several interesting things we can see from the plot. A fast rise of the dot-com bubble is seen in years 1999 and 2000, when there was a large amount of high-tech companies going public (see Ljungqvist and Wilhelm (2003)) with a subsequent burst. Mainly high P/S ratio companies were affected and low ratio, i.e. more traditional companies, were not affected at all. The 2008-2009 market crash, where all quintiles suffered major downturns, is also clearly visible, with a subsequent recovery.

We also ran the same analysis on the European market. We used the following countries: Austria, Belgium, Denmark, Finland, France, Germany, Italy, the Netherlands, Norway, Portugal, Slovenia, Spain, Sweden, Switzerland and the United Kingdom. At each rebalance date we take the 500 largest companies across these countries, divide them into quintiles, and apply the same analysis as we did for the United States.

We can see in Table 3.6 and in Figure 3.2 similarly good results for the European market. This gives us further reassurance that the ratio really has predictive power. Again in this case Volatility and Maximum Drawdown increase with lower quintiles, but Return/Volatility does as well, implying that the risk-adjusted returns improve.

We will do the same for the EV/EBIT ratio for which a simple backtest

Table 3.6: Various statistics for P/S ratio quintile portfolios for the 500 largest companies in Europe. Buy and hold (B&H) is included for comparison.

		Price to sales ratio quintile				
	B&H	Q1	Q2	Q3	Q4	Q5
Returns (arithm. avg.)	8.09%	10.25%	9.80%	6.39%	6.62%	5.58%
Returns (geom. avg.)	6.97%	8.62%	8.51%	5.12%	5.28%	4.28%
Volatility	16.10%	19.56%	17.78%	16.53%	17.02%	16.38%
Max Drawdown	-56.60%	-62.89%	-60.19%	-54.80%	-53.48%	-59.62%
Returns / Volatility	0.50	0.52	0.55	0.39	0.39	0.34
$100 becomes	$433	$604	$590	$296	$306	$249

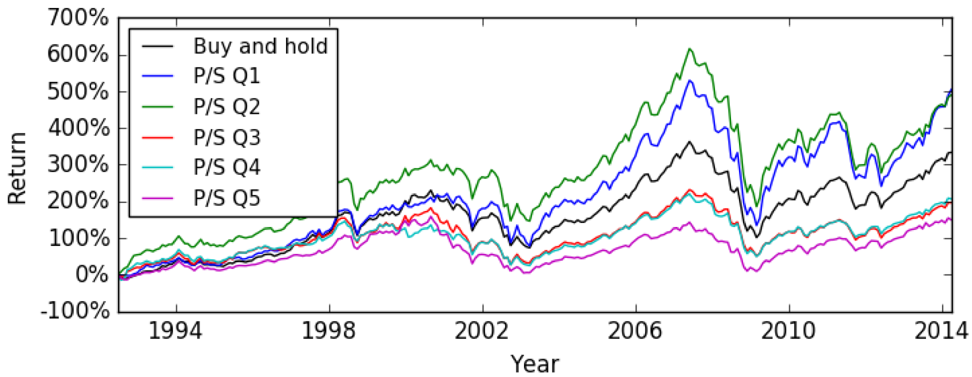

Figure 3.2: Returns for P/S quintiles for the European market. Buy and hold (B&H) is included for comparison.

also worked well. First, we run a backtest on companies from the United States, where we rebalance every 3 months and have access to earnings announcement dates.

The results are shown in Table 3.7 and Figure 3.3. We can again see a progression in returns from higher to lower quintiles. Return/Volatility is also greatly improved with higher quintiles compared to lower quintiles and to buy and hold. Although returns are similar, we would rather choose the EV/EBIT ratio because of the lower volatility in the portfolio, which is what investors usually want.

Table 3.7: Various statistics for EV/EBIT ratio quintile portfolios for the 500 largest companies in the United States. Buy and hold (B&H) is included for comparison.

		EV/EBIT ratio quintile				
	B&H	Q1	Q2	Q3	Q4	Q5
Returns (arithm. avg.)	9.59%	12.92%	11.29%	10.82%	9.87%	9.86%
Returns (geom. avg.)	8.52%	11.97%	10.53%	10.13%	8.71%	7.48%
Volatility	16.46%	17.51%	15.56%	14.86%	17.09%	22.64%
Max Drawdown	-54.87%	-57.53%	-51.24%	-49.06%	-49.59%	-67.15%
Returns / Volatility	0.58	0.74	0.73	0.73	0.58	0.44
$100 becomes	$592	$1,170	$883	$816	$615	$480

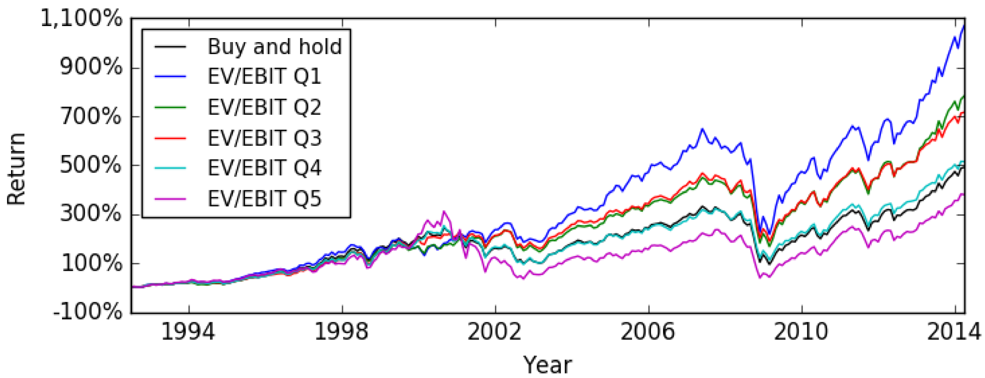

Figure 3.3: Returns for EV/EBIT quintiles for the United States market. Buy and hold (B&H) is included for comparison.

The results for European market are displayed in Table 3.8 and Figure 3.4. We find evidence of the EV/EBIT ratio having great predictive power here as well, which gives us confidence in the ratio. There is a smaller number of companies in the portfolio at each point in time compared to the United States portfolio, since there is a smaller amount of companies that have sufficient data to calculate the EV/EBIT ratio. Therefore, we do not consider these results as reliable.

Table 3.8: Various statistics for EV/EBIT ratio quintile portfolios for the 500 largest companies in Europe. Buy and hold (B&H) is included for comparison.

		EV/EBIT ratio quintile				
	B&H	Q1	Q2	Q3	Q4	Q5
Returns (arithm. avg.)	8.09%	12.25%	9.03%	8.83%	6.94%	4.56%
Returns (geom. avg.)	6.97%	11.24%	8.16%	7.89%	5.68%	2.44%
Volatility	16.10%	17.53%	15.04%	15.39%	16.52%	20.41%
Max Drawdown	-56.60%	-55.35%	-53.70%	-52.26%	-57.89%	-77.29%
Returns / Volatility	0.50	0.70	0.60	0.57	0.42	0.22
$100 becomes	$433	$1,015	$551	$522	$332	$169

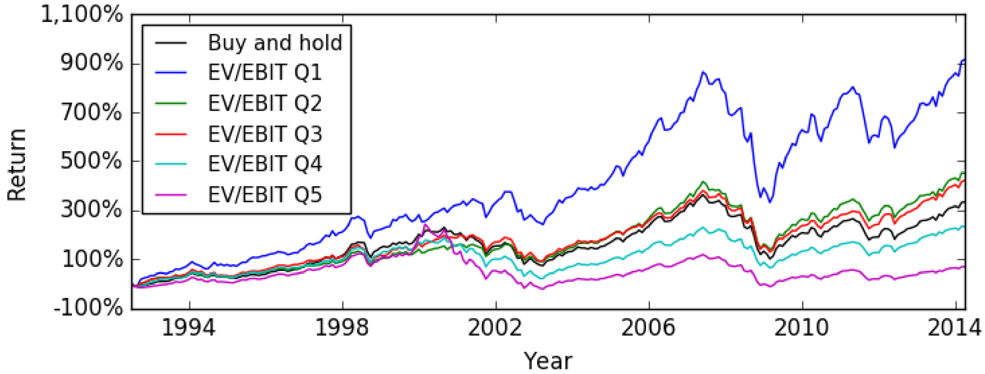

Figure 3.4: Returns for EV/EBIT quintiles for the European market. Buy and hold (B&H) is included for comparison.

Chapter 4

More complex strategies

Now we move on to strategies that are more complex than calculating a single ratio and investing according to it. We have selected three of the papers published on stock selection by the use of accounting data. Each paper has a unique methodology of stock selection.

We will take these three academic papers and implement their strategies to see if we are able to reproduce their results on our dataset.

4.1 Piotroski strategy

Piotroski (2000) attempts to differentiate between value companies that are financially distressed and those that are not. Here it is supposed that the latter are of higher quality and will tend to outperform the former. He develops a test of company quality by assigning a discrete score from 0 to 9 to each company that is considered for inclusion in the portfolio.

The score is called F_SCORE and is a sum of 9 binary signals, each of which is a member of one group. The first group is profitability, the second is leverage, liquidity and source of funds and, the third is operating efficiency.

Profitability:

- return on assets > 0,

- cash flow from operating activities > 0,
- cash flow from operating activities - return on assets > 0,
- current year's return on assets - prior year's return on assets > 0.

Leverage, liquidity and source of funds:

- last year's long term debt - prior year's long term debt > 0 (scaled by total assets),
- last year's current ratio - prior year's current ratio > 0,
- firm did not issue common equity.

Operating efficiency:

- last year's gross margin ratio - prior year's gross margin ratio > 0 (scaled by total assets),
- last year's asset turnover ratio - prior year's asset turnover ratio > 0.

By summing these binary signals, we get a score for each company.

Piotroski (2000) combines multiple accounting statement items into a single numerical score. It is the only paper out of the three we analyze, where the accounting items used are chosen by the author himself, as opposed to a quantitative model. The original data for testing the strategy span from 1976 to 1996. Since the author has seen the data when choosing the accounting items, there was no discretionary in-sample/out-of-sample data selection and the performance testing was done on the same dataset.

We are able to run a backtest 16 years after the publication and test if the model works on our data and after the publication. Piotroski (2000) actually calculated F_SCORE on a universe of companies that was already pre-selected to be undervalued. His aim with F_SCORE was to avoid so-called *value traps* - these are companies that look undervalued but their price continues to fall. We avoid this pre-selection step here to see how F_SCORE behaves in general on all companies, not just on undervalued companies. We will do this step in Chapter 5 where we combine value and quality strategies.

As in the previous chapter, we do a rebalance once a year on July 1. We take annual reports of all the companies that had their fiscal year end between 5 months and 17 months before the rebalance date (for a window of 1 year). F_SCORE is calculated for each annual report, as well as subsequent 1 year return. Average return is calculated for each score separately and the process is repeated on the next rebalance date. Geometric average is then used across all years to calculate the final average excess return for each F_SCORE. We do a backtest from 1993 to 2013 inclusive.

Table 4.1: Geometric average of average annual excess returns for companies with certain F_SCORE in the United States market. There were no companies with F_SCORE 0.

	N	Excess return
F_SCORE 0	0	nan
F_SCORE 1	2	-6.84%
F_SCORE 2	18	-9.43%
F_SCORE 3	107	-14.01%
F_SCORE 4	425	-7.13%
F_SCORE 5	979	0.50%
F_SCORE 6	1381	-1.08%
F_SCORE 7	1323	0.77%
F_SCORE 8	850	2.31%
F_SCORE 9	225	2.76%

We can see the results in Table 4.1. We do not have any cases where F_SCORE is 0 and there are only a few instances of F_SCORE 1 and 2. We therefore consider those results highly unreliable. But it is clear from our backtest that companies with a high F_SCORE tend to have higher excess returns on average than those with low F_SCORE, and do, on average, outperform the market.

To increase the number of instances, we create 3 groups instead of 10: those with F_SCORE of either 7, 8, or 9, those with F_SCORE of 4, 5, or 6 and those with F_SCORE of 0, 1, 2, or 3. We invest into companies of each group and look at how each portfolio performed. The results are displayed in Table 4.2.

Table 4.2: Geometric average of average annual excess returns for F_SCORE groups for the United States market.

	N	Excess return
F_SCORE 0, 1, 2, 3	127	-6.81%
F_SCORE 4, 5, 6	2785	-0.96%
F_SCORE 7, 8, 9	2398	1.45%

Again, there is not a lot of cases where a company has F_SCORE of 0, 1, 2 or 3. We display those results, but do not consider them reliable.

Now that we have established that F_SCORE is able to predict future returns, we will generate an actual portfolio value chart and calculate various additional statistics. We will also try to improve the returns by rebalancing every 3 months.

We are allowed to invest in all companies that have the desired F_SCORE on an annual accounting report that was released in a 12 month window 5 months before the rebalance date. Every 3 months we check whether there are any new annual reports we have access to and add new companies or remove those that do not fit the criteria anymore. Every company gets an equal share of our investment. Here we suppose that stock shares are divisible which means that we are able to buy a real number of shares instead of only an integer. For now, we ignore transaction costs.

In Table 4.3 we can see various statistics for each of the three groups. Note that calculated statistics for group *F_SCORE 0, 1, 2, 3* might not be representative as there are few instances of those scores, see Table 4.1.

Figure 4.1 graphically displays how portfolio value fluctuated through the years.

How about investing only in companies of certain F_SCORE, opposed to having three groups?

Table 4.3: Various statistics for F_SCORE group portfolios for the 500 largest companies in the United States. Buy and hold (B&H) is included for comparison.

		Piotroski F_SCORE		
	B&H	7, 8, 9	4, 5, 6	0, 1, 2, 3
Returns (arithm. avg.)	9.34%	10.60%	7.91%	4.51%
Returns (geom. avg.)	8.18%	9.56%	6.46%	-0.37%
Volatility	16.83%	16.86%	17.86%	31.06%
Max Drawdown	-54.87%	-51.36%	-56.72%	-84.54%
Returns / Volatility	0.56	0.63	0.44	0.15
$100 becomes	$502	$649	$361	$93

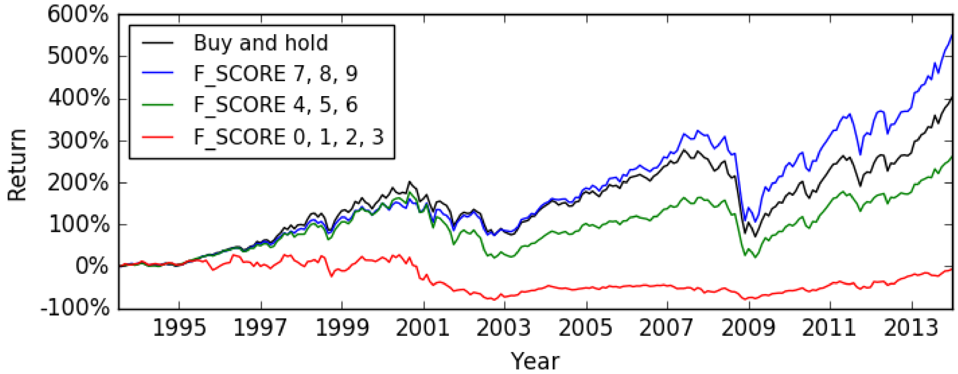

Figure 4.1: Returns for grouped F_SCORE strategies for the United States market. Buy and hold (B&H) is included for comparison.

In Table 4.4 and Figure 4.2 we can see the results how the same investment strategy we used before would work when investing in companies that belong to certain F_SCORE. It is surprising how consistent the results are. Almost every F_SCORE group has higher arithmetic and geometric average returns than the group below.

Note that we have fewer instances in each group than we did before with three groups, so there is a higher chance of skewed results.

Table 4.4: Various statistics for single F_SCORE strategies for the largest 500 companies in United States. Buy and hold (B&H) is included for comparison.

		Piotroski F_SCORE of							
	B&H	9	8	7	6	5	4	3	2
Returns (arithm. avg.)	9.34%	12.77%	11.93%	10.07%	8.27%	8.16%	3.78%	0.07%	1.75%
Returns (geom. avg.)	8.18%	11.49%	10.90%	9.00%	6.95%	6.59%	1.33%	-6.18%	-2.72%
Volatility	16.83%	19.02%	17.36%	16.71%	17.36%	18.47%	21.89%	35.49%	29.73%
Max Drawdown	-54.87%	-50.57%	-50.38%	-53.71%	-49.75%	-59.44%	-80.23%	-94.92%	-91.38%
Returns / Volatility	0.56	0.67	0.69	0.60	0.48	0.44	0.17	0.00	0.06
$100 becomes	$502	$930	$834	$585	$396	$370	$131	$27	$57

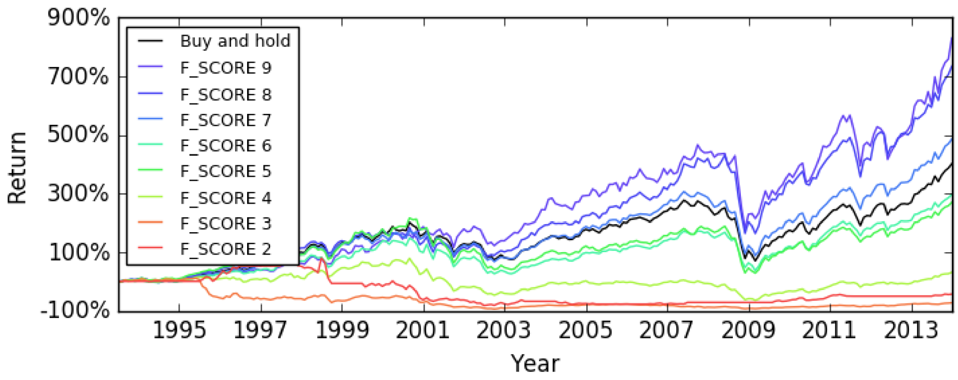

Figure 4.2: Returns for single F_SCORE strategies for the United States market. Buy and hold (B&H) is included for comparison.

4.1.1 The European market

So far our universe of companies has been the largest 500 companies in the United States. We will also try the Piotroski strategy on another market. Our universe will now be the 500 largest companies in Europe.

The results are displayed in Table 4.5 and Figure 4.3. Again, there is only a small number of companies with F_SCORE 0, 1, 2 or 3 and the results are therefore not reliable for this group. We can see however, that overperformance of a group with F_SCORE 7, 8, 9 and some underperformance of a group with F_SCORE 4, 5, 6, as we did with companies from the United States.

4.1.2 Including transaction costs

In reality, however, we have to pay transaction costs when buying and selling stocks. For certain online brokerages, these costs can, at the time of this writing, be as low as 0.05% of the traded volume. However, since transaction costs were higher in the past, we will be more conservative. We will make two simulations - where we suppose we have to pay either 0.5% or 1% of the traded volume.

We will repeat the previous backtest for our best performing group, where

Table 4.5: Various statistics for F_SCORE group portfolios for the 500 largest companies in Europe. Buy and hold (B&H) is included for comparison.

		Piotroski F_SCORE		
	B&H	7, 8 or 9	4, 5 or 6	0, 1, 2 or 3
Returns (arithm. avg.)	7.77%	10.88%	7.41%	6.94%
Returns (geom. avg.)	6.60%	9.81%	5.92%	3.35%
Volatility	16.30%	17.06%	17.88%	27.23%
Max Drawdown	-56.60%	-49.68%	-52.14%	-71.57%
Returns / Volatility	0.48	0.64	0.41	0.26
$100 becomes	$371	$681	$325	$196

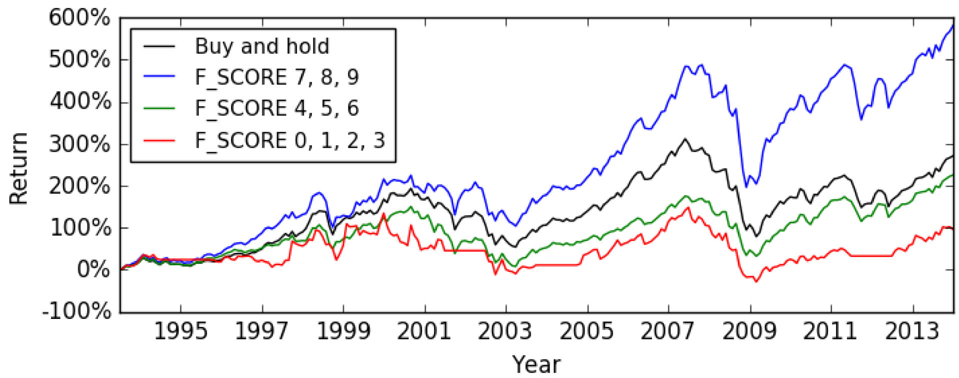

Figure 4.3: Returns for grouped F_SCORE strategies for the European market. Buy and hold (B&H) is included for comparison.

we invested in companies from the universe of the 500 largest US companies and pick only those with F_SCORE of 7, 8, or 9. Our software simulates the transactions costs that occur when buying or selling. A company is bought when the latest available accounting report indicates that the company's F_SCORE is 7, 8, or 9 and sold when it goes below 7. Every three months, if companies are bought or sold, the transaction costs are deducted from the portfolio value.

In Table 4.6 and Figure 4.4 we can see that the strategy still beats buy and hold even when including transaction costs, although with 1% costs the

Table 4.6: Various statistics for portfolios of companies with F_SCORE 7, 8 or 9 for the 500 largest companies in the United States with transaction costs included. Buy and hold (B&H) is included for comparison.

		Transaction costs		
	B&H	0%	0.5%	1%
Returns (arithm. avg.)	9.34%	10.81%	10.17%	9.53%
Returns (geom. avg.)	8.18%	9.81%	9.11%	8.42%
Volatility	16.83%	16.68%	16.69%	16.72%
Max Drawdown	-54.87%	-50.96%	-51.29%	-51.61%
Returns / Volatility	0.56	0.65	0.61	0.57
$100 becomes	$502	$682	$598	$524

overperformance is minimal. We cannot, however, discard the strategy's historical overperformance. As we can see, it was possible to do it even with high transaction costs.

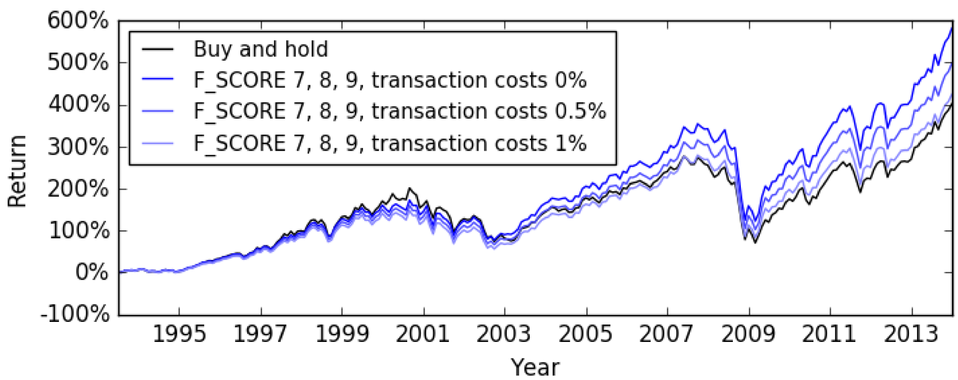

Figure 4.4: Returns for F_SCOREs 7, 8 or 9, with transaction costs included, for the United States market. Buy and hold (B&H) is included for comparison.

4.2 Ou & Penman strategy

We will now look at two papers that utilize statistical models to evaluate the importance of accounting items. Unlike the models we used so far, they do not arbitrarily select accounting items in advance.

Ou and Penman (1989) approach investment decisions by utilizing company's latest annual accounting data to try to predict whether company's next year's earnings, compared to previous year's earnings, will expand or contract more than usual.

The indicator is shown in the following equation:

$$\text{earnings direction} = \begin{cases} 1, & e.p.s._{t+1} - e.p.s._{t} - drift > 0, \\ 0, & \text{otherwise,} \end{cases}$$

where $e.p.s._t$ is earnings without extraordinary items per share in year t and $drift$ is average $e.p.s.$ change in the four years prior to $t+1$. This indicator is a binary value 1 if the company's next year's change in earnings without extraordinary items per share is larger than the trend of the past four years, and 0 otherwise. They suppose this indicator is related to the company's future stock performance.

They use logistic regression as their classification method. In their fundamental data they have 68 accounting items for each company. Not every item is available for every company in every year, however, since there are missing items in the dataset. Their strategy requires that all items exist for training and testing. Because of that, prior to using logistic regression to train the model, they decrease the amount of accounting items by selecting those that have higher prediction ability. By doing that, the number of train and test instances increases, because fewer instances have missing data.

The selection of accounting items is done in the following steps:

1. They evaluate each accounting item's sole ability to predict next year's earnings direction. They run logistic regression on each item separately

and calculate its coefficient, chi-square statistic, and p-value. Only items that have p-value of less than or equal to 0.1 are kept.

2. The remaining items are used, now jointly, in a logistic regression. Again, only items that have p-value of less than or equal to 0.1 are kept. Here, instances with missing values are not disregarded yet.

3. Step-wise regression is used on the remaining items and again, only those with p-value of less than or equal to 0.1 are kept. The authors report keeping most items at this stage in their dataset.

The items that remain are used to train the logistic model and evaluate it on out-of-sample data. Any instance that doesn't have values for complete set of items is omitted.

The reported percentage of successful predictions in their dataset is approximately 66%. But that does not tell us much since this information could already be incorporated the stock prices. That's why they test the viability of the strategy by running a backtest where their predictions of earnings directions are used to hypothetically create long and short positions.

They split their data into three periods: 1965 - 1972, 1973 - 1977 and 1978 - 1983. They

(a) train the model on the first period and evaluate it on out-of-sample on the second period, and

(b) train on the second period and evaluate it on the third period.

Their trading strategy is the following: at the end of the third month, after the end of company's fiscal year, they are allowed to take the accounting data reported for that year (they assume the data have been publicly available at that point) and insert the data into the logistic regression model that was trained on the previous period. The output of the model is a number P which is defined as the probability that the next year's earnings change compared to the trend is positive. They choose to put long positions every time P is higher than 0.6/0.5 and put short positions when P is 0.4/0.5 or less. The positions are kept for 12/24 months after entering and the result is compared to the equal-weighted long portfolio of all the companies.

They analyzed annual accounting data of around 2000 companies each year between 1965 and 1983. We have data for the 500 largest US companies between years 1992 and 2013. We therefore test their investment strategy on data completely unknown to them when the paper was published in 1989.

We have not completely replicated their approach. Instead of separating data into three parts of equal size, where model trained on part 1 is used to predict earnings direction on part 2, and similarly for parts 2 and 3, we use a rolling window approach to train the model on a 10-year window and test it on the subsequent year. That means we first train our model on all the data from 1995 to 2004, inclusive, and test it on year 2005, then slide by one year, up to training on the period from 2003 to 2012 and testing on year 2013. We chose number 10 in advance, to capture the whole business cycle.

The original paper allowed the use of annual accounting data at the end of the third month after fiscal year end. Our dataset also includes earnings dates data, i.e. the dates when the annual accounting reports were released. In our analysis we do both - we take the original approach and the approach with using earnings announcement dates.

4.2.1 Dataset

Ou and Penman (1989) used 68 accounting items as data for their model. Our dataset contains all of the items except *change in production*, *change in total uses of funds*, and *change in total sources of funds*. We do not consider this to be a significant problem since none of the removed accounting items pass the selection in the original test. Also, items *repayment lt per lt* and *issuance lt per lt* do not appear in the earlier part of our dataset. Therefore, we skip them, keeping 63 out of the 68 original accounting items.

Both in training or test part, they allowed only instances without missing data. Having a larger number of accounting items means there is a larger possibility that at least one item value is missing. That means there are fewer instances to train the model on and fewer instances to evaluate it on. That is why we choose to do an additional pre-selection step where we keep only accounting items that are contained in at least 70% of the annual accounting statements.

From now on, we use the same steps for selecting accounting items as the original paper. First, accounting item's sole ability to predict earnings direction is computed. Table A.1 displays, for each accounting item that was used in the model, the number of item's occurrences in the data, coefficient when compared to next year's earnings direction (i.e. the sole prediction ability), and its p-value. The data for the first and the last period out of nine 10-year periods, where we train our nine models, are shown. We display only 54 items out of the original 68, since we discard any item that has more than 30% of its values missing. After this stage, all items that have p-values of their coefficient estimates less than or equal to 0.10 are kept. The original paper keeps 34 items out of 68 (50%) in both periods. In our dataset, there are 18 and 20 items, in the first and the last period, respectively, with p-value of less or equal to 0.1.

Then, the number of accounting items is further reduced by performing a selection of items with p-value 0.1 or lower in a multivariate model, which results in the original paper keeping 19 and 18 items for the first and second period, respectively. In the third stage, items with p-value 0.1 or less when performing stepwise regression are selected. After doing that, the paper ends up with 19 accounting items for period 1 and 15 for period 2. In our dataset, after the multivariate selection, the number of kept items for the two periods drops to 8 and 11. No items in these two periods are dropped after the stepwise regression, although they can be and are dropped in other periods.

4.2.2 Perfect foresight

We will first look at a hypothetical strategy where we have the ability to see what the next year's earnings direction (adjusted by drift) is. Long positions are made for companies with positive earnings direction and short positions for the rest.

Returns of this strategy are compared to the returns of an equal-weight buy and hold portfolio of companies that have the required data. For all companies that have a positive next year's earnings direction, relative returns for companies that are long in the portfolio are thus calculated as $R_{positive-direction-strategy} - R_{buy-and-hold}$ and for shorted companies they are calculated as $R_{negative-direction-strategy} - R_{buy-and-hold}$.

This strategy cannot be used in real life because we do not know what the earnings direction will be. However, by calculating returns for such a strategy, we find out if having the perfect foresight of the earnings direction has an influence on stock returns.

Geometric averages of annual excess returns are displayed in Table 4.7. We can see that the future earnings direction is indeed a valuable information. Having this information improves both long and short strategy. Note that only companies that have a full set of accounting items we require in later realistic models are included here.

Table 4.7: Geometric average of average annual excess returns for a hypothetical strategy where we have information of what next year's direction in earnings compared to current year's adjusted by drift is. Done on the United States market.

Next year's change	N	Excess return
change e.p.s. - trend ≤ 0	1322	-2.66%
change e.p.s. - trend > 0	1406	2.78%

4.2.3 Training a logistic model

Now that we know that having information about future earnings direction can result in a strategy that outperforms buy and hold, we will attempt to create a model that tries to predict the direction. We take the training part of the data and do the following:

1. Each first day of a month we separate newly available yearly accounting statements into two groups - those with positive next year's future earnings direction, and those with negative.

2. A 12-month window, from where we can choose companies, is used, so we avoid seasonal variations on number of companies. We are allowed to look at the accounting statement for a company at the beginning of the fourth month after the end of the fiscal year.

3. Future excess returns, compared to buy and hold, are calculated.

4. Logistic regression is used to model the relationship between accounting items and direction.

4.2.4 Testing the logistic model

We can now say that each accounting statement and earnings direction represents a row in a matrix. We use logistic regression to train our model and test it on a subsequent year. Because we use 10-year rolling window, we are able to test years 2005 to 2013. We take long positions in cases where P, i.e. a real number from 0 to 1 our model returned on a new test instance (accounting statement), is larger than 0.6, and short position when P is less than or equal to 0.4.

Table 4.8: Geometric average of average annual excess returns for companies with certain P for the United States market.

Model forecast	N	Excess return
$P \leq 0.4$	458	-0.40%
$P > 0.6$	391	-2.56%

In Table 4.8 we can see the results. This is not what we expected since the return should increase with companies for which our model predicts the probability of increased earnings is high.

Now, we will split our groups into five parts for a closer look at how different P values affected returns.

Table 4.9: Geometric average of average annual excess returns for companies with certain P for the United States market.

Model forecast	N	Excess return
$P \leq 0.2$	47	-0.74%
$0.2 < P \leq 0.4$	410	-0.28%
$0.4 < P \leq 0.6$	1760	0.45%
$0.6 < P \leq 0.8$	350	-2.21%
$P > 0.8$	41	-2.83%

In Table 4.9 we can see that both lowest and highest P values do underperform and the middle part overperforms, which is not what we expected.

Now we will do as we did with ratios - we will split companies into approximately 5 equally sized groups and test each group's relative performance.

Table 4.10: Geometric average of average annual excess returns for companies belonging to a certain quintile according to P for the United States market.

Quintile	N	Excess return
Q1	517	-0.94%
Q2	521	1.57%
Q3	523	-0.02%
Q4	521	0.14%
Q5	526	-1.18%

In Table 4.10 we can see average excess returns of the strategy. Companies in the lowest quintile had a lower return than the market, which is what we would expect. However, the highest quintile also underperformed and this is contrary to expectations.

4.2.5 Incorporating earnings dates

Since we have access to earnings dates, i.e. dates when companies publicly released their annual accounting data, we also use those instead of arbitrarily choosing how many months to delay. Our analysis showed that, while companies' earnings dates of the 500 largest companies in the United States are on average around 33 days after the fiscal year end, the number varies a lot. We therefore use this additional data to more precisely determine at what point in our backtest are we allowed to use certain accounting data.

In Figure 4.5 we can see the distribution of number of days between annual fiscal year end date and subsequent earnings date for the 500 largest companies in the United States, from 1993 to 2014. The majority of earnings are reported before 60 days after the fiscal year end and the average is 33 days.

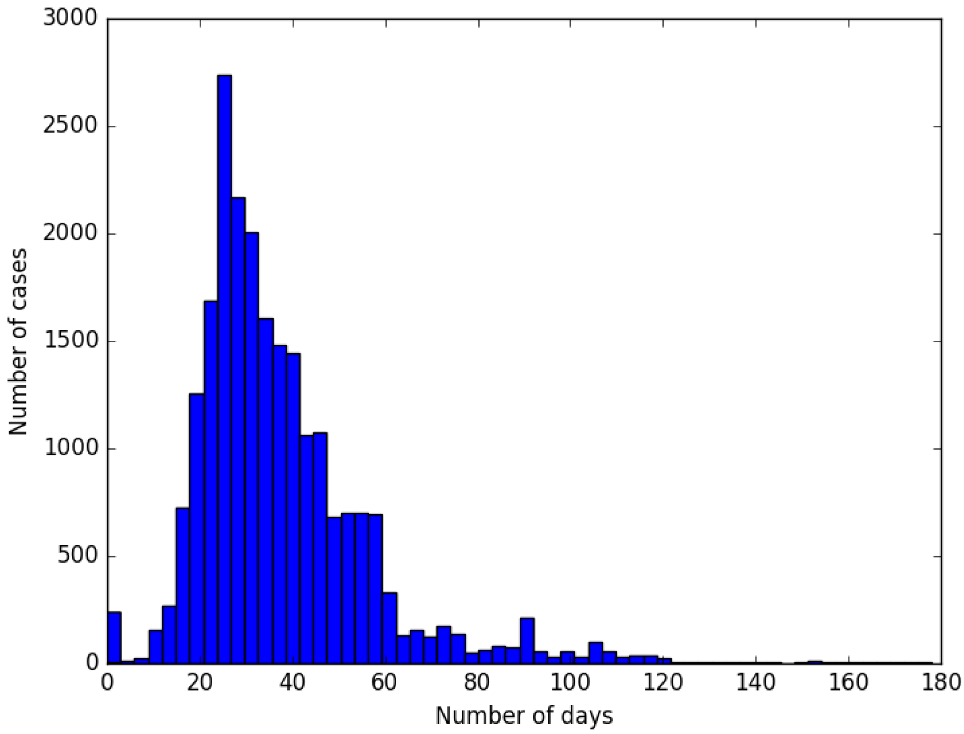

Figure 4.5: Frequency of number of days between annual fiscal year end date and subsequent earnings date for the 500 largest companies in the United States.

In order to be more responsive, we will now do 4 rebalances in a year, one every 3 months. Our investment strategy will otherwise stay the same.

For some companies, we allow access to accounting items sooner than in the previous approach, and also remove the possibility of seeing into the future in cases where it took companies longer to release their information.

Table 4.11: Geometric average of average annual excess returns for companies belonging to a certain quintile according to P for the United States market. Earnings announcement dates are here used.

Quintile	N	Excess return
Q1	495	0.46%
Q2	509	-0.03%
Q3	508	-1.00%
Q4	509	-0.19%
Q5	522	0.39%

The results are shown in Table 4.11.

In this case, we again fail to get results we would expect according to the assumption of the strategy. We also fail to achieve consistent excess returns with doing rebalances every 3 months going back to allowing access to accounting statements with the end of fiscal year between 4 and 5 months prior to rebalance date or when combining the two approaches by using earnings announcement dates when available, otherwise allowing access between 4 and 5 months prior to rebalance date.

Using a strategy that rebalances every 3 months with a 3 month accounting statement window is problematic, because in certain periods there is not enough companies that ended their fiscal year, to do a meaningful separation. That is why we have also tried using a 12 month accounting statement window and using earnings dates. That means that the window now overlaps, but there is more choice of companies and the number of companies in each group grows. Here we have also failed to get the results of the original paper.

4.2.6 The European market

Repeating the strategy on the European market yields similar results. In Table 4.12 results for two portfolios are displayed. We would expect the excess return for P ≤ 0.4 to be negative and the excess return for P > 0.6 to be positive. The results do not match the expectations.

In Table 4.13, results for portfolios split into quintiles are shown. Again, we do not see convincing results where lower quintiles underperform. The model forecasts negative earnings direction for lower quintiles and positive earnings direction for higher quintiles and that is not reflected in the returns.

The results do not confirm our expectations.

Table 4.12: Geometric average of average annual excess returns for companies with certain P for the European market.

Model forecast	N	Excess return
P ≤ 0.4	113	2.84%
P > 0.6	344	0.36%

Table 4.13: Geometric average of average annual excess returns for companies belonging to a certain quintile according to P for the European market. Earnings announcement dates are used here.

Quintile	N	Excess return
Q1	245	2.30%
Q2	248	-0.76%
Q3	251	-1.55%
Q4	248	-0.03%
Q5	253	-0.63%

4.2.7 Possible reasons for different results

We cannot compare our results directly with the results from the original paper since our investment strategy is a bit different. We aim to be consistent with other strategies in this thesis so we can compare them. The original paper calculated simple arithmetic average of all the 12-month returns for a long and short portfolio separately, and when combining them, got an excess return of 11.52%. These are clearly not the results we see here.

There are several possible reasons why we haven't achieved results compared to the original paper. The reason for this might be in the universe itself. We only have access to the 500 largest companies. The original approach uses a much larger universe, from 2500 to 3000 companies. There's a possibility that the strategy works on smaller companies.

Another reason might be that the strategy actually stopped working. The paper was published in 1989 and our dataset starts at 1992. We also chose to have a 10-year window on which we train our model on. Since calculating drift, which we use in the earnings direction, requires a few years of data by itself, we can only start testing the strategy from 2005 onwards. Up to 2013 that mean 9 years of data, which might not be enough to sufficiently validate or reject a strategy.

Another reason might be that the quality of data is different. With a complete set of data, where there are no missing accounting items, the model might work better. Because of missing items there is a larger possibility of distorted results, even if the model is truly one that would give us a set of companies that will overperform, because the amount of instances on which we are testing the model is smaller.

This is a problem for all the strategies that use a large amount of accounting data. The larger the set of accounting items we require, the bigger the possibility at least one item is missing. Simpler strategies such as simple value investing with few required accounting items, and even Piotroski F_SCORE strategy, gives us much more instances on which to test.

4.3 Bartram strategy

Bartram and Grinblatt (2015) is another paper we try to replicate that uses statistical techniques to predict stock market returns. The paper starts with an assumption that, on average, companies are fairly valued. That is, that the price of a market as a whole is where it should intrinsically be, but that subsections of the market might be undervalued or overvalued. They used quarterly accounting data to model market capitalization. Their dataset spans years from 1976 to 2012 and 28 most common accounting items are used.

The model is trained using linear regression on data where each row represents a company - month pair. Each row contains 14 balance sheet items from the latest quarterly accounting statement and 14 income statement items that are sums of items from the last 4 quarterly accounting statement. This summation is done in order to avoid seasonal variations. The dependent variable is market capitalization, which is what we are trying to predict in the test part of the dataset.

Latest accounting statement data and market capitalization are recorded for each month in the training part of the dataset. Even though the accounting data stays the same for three months at a time, since quarterly accounting statement is released 4 times a year, the market capitalization changes more frequently. That is because of the changes in company price and because of the changes in the number of shares outstanding, in case company chooses to issue new shares or if options to get shares are executed.

The model is then tested on out-of-sample data. The linear combination of accounting items for a company j in a month t, with weights that were calculated in the training part, gives a market capitalization prediction $P_t(j)$. Actual market capitalization $V_t(j)$ is used to calculate the degree of mispricing $M_t(j)$. The mispricing signal is calculated as:

$$M_t(j) = \frac{P_t(j) - V_t(j)}{V_t(j)}.$$

The hypothesis being tested in the paper is that stocks with highly positive mispricing signal will in the future overperform, and that, conversely, stocks with low or negative mispricing will underperform. Companies are on each rabalancing date split into 5 quintiles where we denote the most overpriced quintile with Q1 and the most underpriced quintile with Q5. Q5 should therefore overperform Q1 in the future.

When testing the Bartram and Grinblatt (2015) strategy, we use the same data as in previous analyses, that is, data of a period from 1995 to 2013, inclusive, whereas the original paper used data from 1977 to 2012. Our analysis also limits the company universe to the 500 largest US companies. We do that, because these companies are more likely to have less missing accounting data and it is more likely that accounting data that exists, are correct.

We implement the strategy by training a linear model on a 10-year window of data, similar to the previous strategy. Each training instance contains balance sheet items from the latest available quarterly accounting statement and income statement items, where each item is a sum of values from 4 latest quarterly accounting statements. We allow the data to be accessed a day after the earnings release date.

For accounting statement data to be included in the training model or considered in testing, we require that all 28 accounting items exist. If there are any missing values among the items in a certain accounting statement, we simply omit the company in that time period.

The investment strategy is similar to what we did before. Once a year, on July 1, we calculate average return for each quintile that is based on M_t, and subtract average return across all companies in that year to get average excess return. At least 5 months and at most 17 months must have passed since particular company's fiscal year end for inclusion. We do that for all the years and calculate geometric average of the average annual excess returns.

In Table 4.14 we can see the results. Lower quintiles do underperform and higher quintiles do overperform, as we expected. The results are much better than the results when reproducing the work of Ou and Penman (1989).

Now we will create and analyze a strategy that is more adaptive. First, we change the rebalancing frequency to 4 times a year instead of 1 time. We

Table 4.14: Geometric average of average annual excess returns for companies in a certain Bartram quintile for the 500 largest companies in the United States.

Quintile	N	Excess return
Q1	604	-1.10%
Q2	608	-0.82%
Q3	606	-0.32%
Q4	608	1.63%
Q5	610	0.27%

Table 4.15: Various statistics for Bartram quintile portfolios for the 500 largest companies in the United States. Buy and hold (B&H) is included for comparison.

		Bartram quintile				
	B&H	Q5	Q4	Q3	Q2	Q1
Returns (arithm. avg.)	8.36%	12.59%	13.22%	8.70%	9.75%	9.06%
Returns (geom. avg.)	6.69%	9.40%	11.47%	7.22%	8.49%	7.80%
Volatility	18.98%	26.42%	21.16%	18.23%	17.45%	17.23%
Max Drawdown	-54.87%	-63.95%	-52.48%	-45.31%	-47.69%	-49.52%
Returns/Volatility	0.44	0.48	0.62	0.48	0.56	0.53
$100 becomes	$172	$213	$249	$180	$199	$188

also use the earnings announcement dates data to more precisely determine when accounting data were made available. We allow using earnings data that were released up to 12 months in the past. That allows us to have a large number of companies in the portfolio at any given time.

We display detailed statistics and graphical results of the strategy in Table 4.15 and Figure 4.6.

Interestingly, each quintile performs better than buy and hold. How is it possible that each group performs better than the universe itself? Since we discard a company on a certain rebalance date if any of the accounting items we require is missing on the accounting statement, not all companies from the universe are included in the quintiles. The companies that pass the

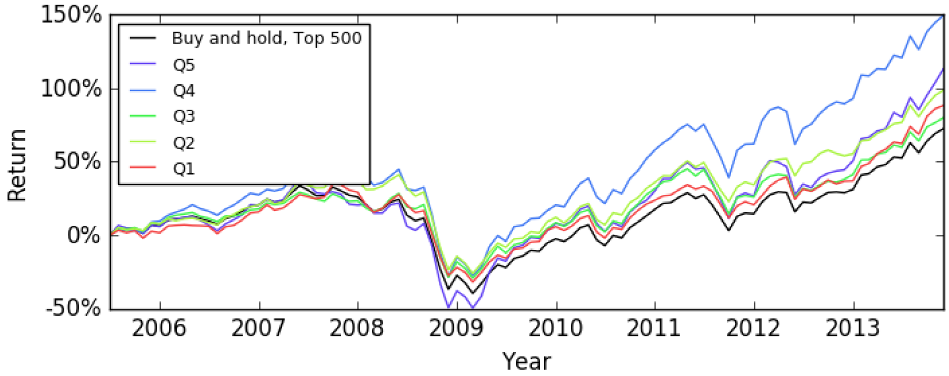

Figure 4.6: Returns for Bartram quintiles for the United States market. Buy and hold (B&H) is included for comparison.

selection apparently overperform the market.

There are two possible explanations as to why this happens. The first reason for overperformance of every quintile might be that there is a look-ahead bias here since the companies that are more important at the present time are less likely to contain missing data because of an error when collecting or aggregating data. The second reason for this might be that companies did actually have some accounting items missing when they reported their accounting data and that such companies do actually tend to underperform the market. We have no way of determining what the real reason is.

However, when comparing higher quintiles to lower quintiles we see that the former achieve slightly better returns, although the results are not as consistent as we would hope.

Results of the strategy on the European market are displayed in Table 4.16 and Figure 4.7. There are again, no clear results, although the lowest quintile performed the worst - that is what we expected.

Table 4.16: Various statistics for Bartram quintile portfolios for the 500 largest companies in Europe. Buy and hold (B&H) is included for comparison.

		Bartram quintile				
	B&H	Q5	Q4	Q3	Q2	Q1
Returns (arithm. avg.)	5.53%	8.36%	8.30%	8.04%	7.28%	4.26%
Returns (geom. avg.)	4.08%	6.27%	6.69%	7.01%	6.43%	2.51%
Volatility	17.21%	20.98%	18.73%	15.62%	14.16%	18.63%
Max Drawdown	-56.60%	-56.69%	-51.87%	-47.42%	-36.66%	-58.51%
Returns/Volatility	0.32	0.40	0.44	0.51	0.51	0.23
$100 becomes	$140	$167	$172	$177	$169	$123

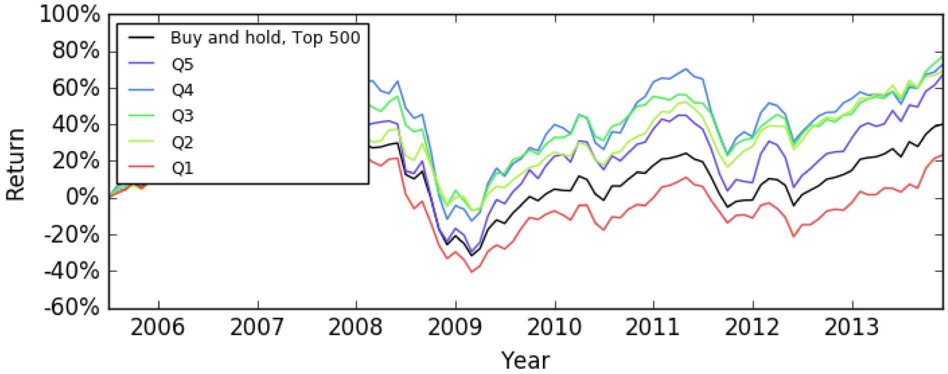

Figure 4.7: Returns for Bartram quintiles for the European market. Buy and hold (B&H) is included for comparison.

Chapter 5

Combining simple strategies

Since simple strategies appear to work better, we will now try to combine them. We will take a value ratio and combine it with Piotroski's F_SCORE. By doing that, we hope to find companies that are undervalued but are at the same time of high quality.

As a value ratio we take the P/S ratio as we have seen that it is a good predictor of future returns. We take the two of most undervalued quintiles and select only companies that have an F_SCORE of 7, 8, or 9.

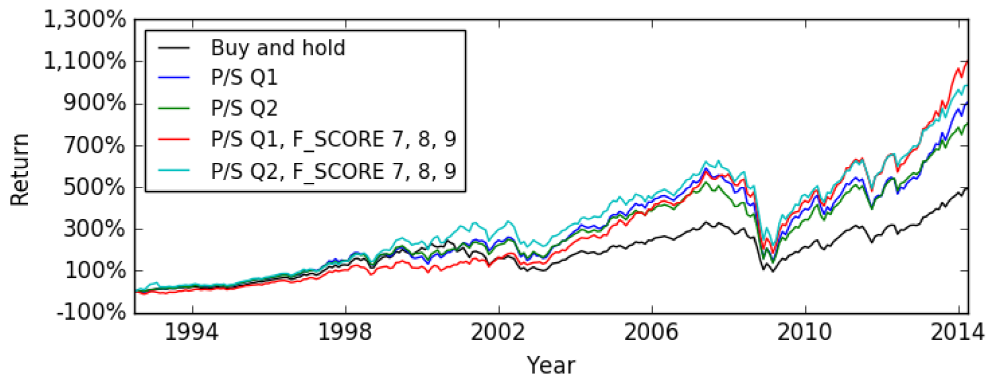

Figure 5.1: Returns for the combination of P/S and F_SCORE for the United States market. Buy and hold (B&H) is included for comparison.

The results are shown in Table 5.1 and Figure 5.1. The returns for

Table 5.1: Various statistics for combining P/S ratio's lowest quintiles and highest F_SCORE portfolios for the 500 largest companies in the United States. Buy and hold (B&H) is included for comparison.

				F_SCORE 7, 8, 9	
	B&H	P/S Q1	P/S Q2	P/S Q1	P/S Q2
Returns (arithm. avg.)	9.59%	12.56%	11.73%	13.17%	12.59%
Returns (geom. avg.)	8.52%	11.19%	10.65%	12.11%	11.59%
Volatility	16.46%	19.26%	17.57%	18.18%	17.65%
Max Drawdown	-54.87%	-65.04%	-61.95%	-58.02%	-56.01%
Returns/Volatility	0.58	0.65	0.67	0.72	0.71
$100 becomes	$592	$1,005	$903	$1,202	$1,085

Table 5.2: Various statistics for combining P/S ratio's lowest quintiles and highest F_SCORE portfolios for the 500 largest companies in Europe. Buy and hold (B&H) is included for comparison.

				F_SCORE 7, 8, 9	
	B&H	P/S Q1	P/S Q2	P/S Q1	P/S Q2
Returns (arithm. avg.)	8.09%	10.25%	9.80%	10.02%	11.65%
Returns (geom. avg.)	6.97%	8.62%	8.51%	8.19%	10.25%
Volatility	16.10%	19.56%	17.78%	20.40%	19.14%
Max Drawdown	-56.60%	-62.89%	-60.19%	-57.45%	-57.05%
Returns/Volatility	0.50	0.52	0.55	0.49	0.61
$100 becomes	$433	$604	$590	$554	$836

both quintiles are improved when allowing only high quality companies. Returns/Volatility is increased as well and maximum drawdown is reduced. We therefore consider this strategy one of the best in the thesis.

We can see the results of the same strategy on the European market in Table 5.2 and Figure 5.2. Here, maximum drawdown is significantly improved in both cases although average return is slightly lower for Q1.

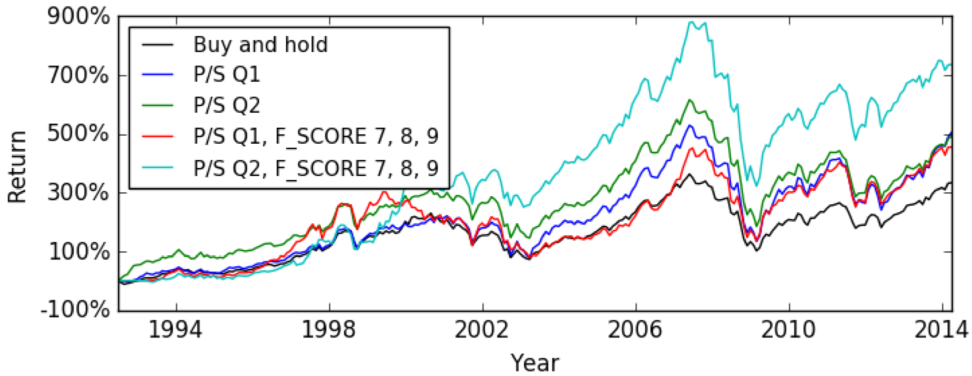

Figure 5.2: Returns for the combination of P/S and F_SCORE for the European market. Buy and hold (B&H) is included for comparison.

Chapter 6

Conclusion

In this thesis we have discovered that simple investment strategies perform better than more complex ones. Simple value investing and quality investing strategies we analyzed achieved much better results than strategies using statistical models on our dataset, which includes the 500 largest companies from the United States and Europe, separately.

We see two possible reasons for the worse performance of more complex strategies. The first is that no qualitative choices were done before doing statistical modeling. The simple strategies make intuitive sense and accounting items used can be explained by economic logic. With the strategies utilizing statistical models on a large amount of accounting items, which means large amount of columns in a matrix, there is always a chance that an otherwise unimportant item is statistically significant in the train part but not afterward.

The second reason is that when requiring a large amount of data, we increase the possibility of missing values. We suspect that a more robust dataset would improve the results of a strategy.

We conclude that finding a strategy that achieves abnormal returns compared to buy and hold is possible for an average investor even when accounting for conservatively large transaction costs and limiting oneself only to largest companies.

6.1 Future work

Our strategies always require a full set of data. If the accounting statement for a company in a particular year has a missing value, the statement is not used in train or test part. There is a possibility of improving this by using, for example, last available value or an average of past values. A learning model that supports missing values could also be used.

Our choice of the dates we do rebalancing on was quite arbitrary. It is possible that more suitable dates exist. We have chosen not to search for better dates in an attempt to avoid optimizing the parameters too much although there might be an opportunity here. We also limit ourselves to doing rebalancing on the first day of a month. A better approach might be to rebalance as soon as accounting data are released. There is a possibility that a large amount of excess returns is already materialized before we even consider certain accounting data.

The models can in future work be improved by the use of regularization. We have not used it because the original papers did not but we suspect regularization could be valuable.

We have limited ourselves to the largest 500 companies in the United States and Europe. An analysis of a larger number of companies could be beneficial. Also, separating companies by sectors or industries might improve overall results.

Bibliography

Alexakis, C., Patra, T., Poshakwale, S., 2010. Predictability of stock returns using financial statement information: evidence on semi-strong efficiency of emerging greek stock market. Applied Financial Economics 20 (16), 1321–1326.

Bartram, S. M., Grinblatt, M., 2015. Fundamental Analysis Works. Available at SSRN 2479817.

Basu, S., 1977. Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. The journal of Finance 32 (3), 663–682.

Bloomberg, 2016. Bloomberg Professional. [Online]. Available at: Subscription Service.

Bolognesi, E., Torluccio, G., Zuccheri, A., 2013. A comparison between capitalization-weighted and equally weighted indexes in the european equity market. Journal of Asset Management 14 (1), 14–26.

Bradshaw, M. T., Richardson, S. A., Sloan, R. G., 2006. The relation between corporate financing activities, analysts' forecasts and stock returns. Journal of Accounting and Economics 42 (1), 53–85.

Faber, M. T., 2007. A quantitative approach to tactical asset allocation. The Journal of Wealth Management, Spring.

Gruber, M. J., 1996. Another puzzle: The growth in actively managed mutual funds. The journal of finance 51 (3), 783–810.

Hunter, J. D., 2007. Matplotlib: A 2d graphics environment. Computing In Science & Engineering 9 (3), 90–95.

Jensen, G. R., ++Johnson, R. R., Mercer, J. M., 1997. New evidence on size and price-to-book effects in stock returns. Financial Analysts Journal 53 (6), 34–42.

Ljungqvist, A., Wilhelm, W. J., 2003. Ipo pricing in the dot-com bubble. The Journal of Finance 58 (2), 723–752.

Malkiel, B. G., Fama, E. F., 1970. Efficient capital markets: A review of theory and empirical work. The journal of Finance 25 (2), 383–417.

Mohanram, P. S., 2005. Separating winners from losers among lowbook-to-market stocks using financial statement analysis. Review of Accounting Studies 10 (2-3), 133–170.

Ou, J. A., Penman, S. H., 1989. Financial statement analysis and the prediction of stock returns. Journal of accounting and economics 11 (4), 295–329.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine learning in Python. Journal of Machine Learning Research 12, 2825–2830.

Piotroski, J. D., 2000. Value investing: The use of historical financial statement information to separate winners from losers. Journal of Accounting Research 38, 1–41.

Seabold, S., Perktold, J., 2010. Statsmodels: Econometric and statistical modeling with python. In: 9th Python in Science Conference.

Setiono, B., Strong, N., 1998. Predicting stock returns using financial statement information. Journal of Business Finance & Accounting 25 (5-6), 631–657.

Appendix A

Ou & Penman single item prediction

Table A.1: Accounting items with their coefficients and p-values for correlation with future earnings direction (adjusted by drift) for two periods for the 500 largest companies in the United States. If an item starts with *change*, that means it is calculated as current year's value minus previous year's value.

		Years 1995 - 2004			Years 2003 - 2012		
#	Item	N	coef.	p-value	N	coef.	p-value
1	change sales to inventory	2439	0.004	0.002	2720	0.000	0.362
2	net profit margin	3142	-1.025	0.000	3618	-1.058	0.000
3	gross margin ratio	2527	0.065	0.448	3040	0.213	0.006
4	days in accs receivable	2874	0.001	0.009	3396	0.002	0.000
5	change working capital	2702	-0.000	0.317	3153	-0.000	0.629
6	change operating income to total assets	3138	0.000	0.769	3614	0.000	0.226
7	change equity to fixed assets	3050	0.020	0.143	3501	0.024	0.049
8	change nopercent return on opening equity	3125	-0.142	0.029	3599	0.004	0.765
9	current ratio	2703	0.020	0.160	3153	0.044	0.004
10	change net profit margin	3142	-0.000	0.638	3618	0.000	0.958
11	change capital expenditure div prev total assets	3016	0.000	0.511	3490	0.000	0.525
12	change working capital to total assets	2702	-0.000	0.398	3153	-0.000	0.602
13	change debt equity ratio	2678	-0.000	0.949	3135	0.000	0.854
14	quick ratio	2695	0.014	0.338	3136	0.047	0.010
15	change inventory turnover	2298	0.002	0.057	2651	0.002	0.086
16	change gross margin ratio	2510	0.000	0.851	3031	0.000	0.504
17	equity to fixed assets	3064	0.012	0.014	3522	0.005	0.111
18	change sales to working capital	2701	-0.000	0.846	3153	-0.000	0.256

19	pretax income to sales	3115	-0.834	0.022	3610	-0.723	0.038
20	times interest earned	2478	-0.000	0.466	2852	0.000	0.953
21	cash flow to total debt	2701	0.273	0.048	3153	0.200	0.089
22	operating income to total assets	3139	-0.210	0.430	3615	-0.208	0.407
23	sales to inventory	2450	-0.000	0.616	2731	0.001	0.077
24	change times interest earned	2434	-0.000	0.362	2807	0.000	0.427
25	purchase treasury stock per stock	3116	3.251	0.001	3562	1.872	0.001
26	change funds	3094	-0.000	0.414	3564	-0.000	0.374
27	sales to fixed assets	3064	0.006	0.029	3522	0.001	0.386
28	change current ratio	2703	-0.000	0.952	3153	-0.001	0.600
29	change lt debt	2751	-0.000	0.386	3259	-0.000	0.513
30	sales to total cash	3134	-0.000	0.683	3611	0.000	0.778
31	working capital to total assets	2703	0.370	0.022	3153	0.475	0.002
32	inventory div total assets	2695	0.508	0.048	3136	0.643	0.012
33	change sales to total assets	3141	0.003	0.114	3616	-0.002	0.185
34	change pretax income to sales	3090	0.000	0.943	3601	0.000	0.228
35	change lt debt to equity	2730	-0.000	0.870	3241	-0.000	0.497
36	sales to total assets	3141	0.046	0.148	3616	0.067	0.028
37	debt equity ratio	2688	-0.000	0.991	3145	0.000	0.971
38	change capital expenditure div total assets	3015	-0.001	0.297	3494	-0.005	0.000
39	return on closing equity	3127	-0.275	0.001	3608	0.002	0.886
40	change nopercent dividend per share	2990	-0.300	0.170	3605	-0.225	0.009
41	lt debt to equity	2778	0.008	0.363	3278	0.001	0.862
42	change sales	3142	-0.002	0.112	3618	-0.001	0.195
43	change total assets	3141	-0.001	0.038	3616	0.001	0.245
44	change inventory	2453	-0.002	0.014	2736	-0.000	0.515
45	net income per cash flows	3100	-0.001	0.530	3580	-0.001	0.503
46	change inventory div total assets	2453	-0.003	0.034	2736	-0.001	0.370
47	sales to accs receivable	2808	0.002	0.123	3308	0.003	0.019
48	change days in accs receivable	2816	0.000	0.738	3315	0.002	0.061
49	change quick ratio	2693	0.001	0.248	3122	0.001	0.149
50	cash divident per cash flows	3077	0.004	0.376	3548	0.003	0.685
51	sales to working capital	2702	0.000	0.912	3153	-0.000	0.132
52	rota	3141	-3.159	0.000	3616	-2.659	0.000
53	inventory turnover	2321	-0.003	0.123	2668	0.003	0.093
54	return on opening equity	3131	-0.329	0.000	3608	-0.031	0.214